

Raw Sequence Listing Error Summary

01PE

ERROR DETECTED

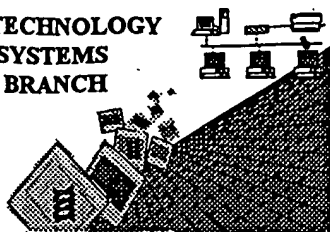
SUGGESTED CORRECTION

SERIAL NUMBER: 10/023,831

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 ✓ Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/023,831
Source: OIPE
Date Processed by STIC: 4/16/02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS-Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202



OIPE

**Does Not Comply
Corrected Diskette Needed**

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/023,831

DATE: 04/16/2002
TIME: 14:35:38

Input Set : A:\EP.txt
Output Set: N:\CRF3\04162002\J023831.raw

3 <110> APPLICANT: VAUGHAN, Paul R.
4 GALANIS, Maria
5 RAMSHAW, John A.M.
6 WERKMEISTER, Jerome A.
8 <120> TITLE OF INVENTION: METHOD FOR PRODUCING, IN YEAST, A HYDROXYLATED TRIPLE
HELICAL PROTEIN,
9 AND YEAST HOST CELLS USEFUL IN SAID METHOD
11 <130> FILE REFERENCE: Q67867
13 <140> CURRENT APPLICATION NUMBER: 10/023,831
14 <141> CURRENT FILING DATE: 2001-12-21
16 <150> PRIOR APPLICATION NUMBER: US 09/297,269
17 <151> PRIOR FILING DATE: 1999-04-28
19 <150> PRIOR APPLICATION NUMBER: PO3310
20 <151> PRIOR FILING DATE: 1996-10-29
22 <150> PRIOR APPLICATION NUMBER: PO4306
23 <151> PRIOR FILING DATE: 1996-12-19
25 <150> PRIOR APPLICATION NUMBER: PCT/AU97/00721
26 <151> PRIOR FILING DATE: 1997-10-29
28 <160> NUMBER OF SEQ ID NOS: 52
30 <170> SOFTWARE: PatentIn version 3.1
32 <210> SEQ ID NO: 1
33 <211> LENGTH: 26
34 <212> TYPE: DNA
35 <213> ORGANISM: Synthetic Construct
37 <400> SEQUENCE: 1
38 ctgtagagga tccccgggta cggagc 26
41 <210> SEQ ID NO: 2
42 <211> LENGTH: 23
43 <212> TYPE: DNA
44 <213> ORGANISM: Synthetic Construct
46 <400> SEQUENCE: 2
47 ttatattgaa ttctcaaaaa ttc 23
50 <210> SEQ ID NO: 3
51 <211> LENGTH: 32
52 <212> TYPE: DNA
53 <213> ORGANISM: Synthetic Construct
55 <400> SEQUENCE: 3
56 tgtaaaatta aagatccca aagatgtggt at 32
59 <210> SEQ ID NO: 4
60 <211> LENGTH: 30
61 <212> TYPE: DNA
62 <213> ORGANISM: Synthetic Construct
64 <400> SEQUENCE: 4
65 gccgggatcc tgtcattcca atgacaacgt 30

The type of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.

↑
- invalid response, see error
summary sheet, item 10

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/023,831

DATE: 04/16/2002

TIME: 14:35:38

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J023831.raw

```

68 <210> SEQ ID NO: 5
69 <211> LENGTH: 44
70 <212> TYPE: DNA
71 <213> ORGANISM: Synthetic Construct
73 <400> SEQUENCE: 5
74 actggacgga tcccgagcgc cccgcctgct ccgtgtccga catg      44
77 <210> SEQ ID NO: 6
78 <211> LENGTH: 24
79 <212> TYPE: DNA
80 <213> ORGANISM: Synthetic Construct
82 <400> SEQUENCE: 6
83 ggttctcctt ggtgacctcc cctt      24
86 <210> SEQ ID NO: 7
87 <211> LENGTH: 24
88 <212> TYPE: DNA
89 <213> ORGANISM: Synthetic Construct
91 <400> SEQUENCE: 7
92 gaaggggagg tcaccaagga gaac      24
95 <210> SEQ ID NO: 8
96 <211> LENGTH: 36
97 <212> TYPE: DNA
98 <213> ORGANISM: Synthetic Construct
100 <400> SEQUENCE: 8
101 ccttcaggat cctattagac ttcattcttc aacagc      36
104 <210> SEQ ID NO: 9
105 <211> LENGTH: 44
106 <212> TYPE: DNA
107 <213> ORGANISM: Synthetic Construct
109 <400> SEQUENCE: 9
110 actggacgga tcccgagcgc cccgcctgct ccgtctccga catg      44
113 <210> SEQ ID NO: 10
114 <211> LENGTH: 35
115 <212> TYPE: DNA
116 <213> ORGANISM: Synthetic Construct
118 <400> SEQUENCE: 10
119 ccttcaggat cctattagac ttcattcttc acagc      35
122 <210> SEQ ID NO: 11
123 <211> LENGTH: 39
124 <212> TYPE: DNA
125 <213> ORGANISM: Synthetic Construct
127 <400> SEQUENCE: 11
128 ctagttgaat tctacacaat gctgcgccgc gctctgctg      39
131 <210> SEQ ID NO: 12
132 <211> LENGTH: 36
133 <212> TYPE: DNA
134 <213> ORGANISM: Synthetic Construct
136 <400> SEQUENCE: 12
137 gcaatggaat tcttattaca gttcgtgcac agcttt      36
140 <210> SEQ ID NO: 13

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/023,831

DATE: 04/16/2002

TIME: 14:35:38

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J023831.raw

141 <211> LENGTH: 4
 142 <212> TYPE: PRT
 143 <213> ORGANISM: Yeast Endoplasmic Reticulum Retention Signal
 145 <400> SEQUENCE: 13
 147 His Asp Glu Leu
 148 1
 151 <210> SEQ ID NO: 14
 152 <211> LENGTH: 5
 153 <212> TYPE: PRT
 154 <213> ORGANISM: Mammalian Endoplasmic Reticulum Retention Signal
 156 <400> SEQUENCE: 14
 158 Lys Asp Ala Glu Leu
 159 1 5
 162 <210> SEQ ID NO: 15
 163 <211> LENGTH: 42
 164 <212> TYPE: DNA
 165 <213> ORGANISM: Synthetic Construct
 167 <400> SEQUENCE: 15
 168 gctagcaagc ttggagctcc aggcccactt gggattgctg gg 42
 171 <210> SEQ ID NO: 16
 172 <211> LENGTH: 45
 173 <212> TYPE: DNA
 174 <213> ORGANISM: Synthetic Construct
 176 <400> SEQUENCE: 16
 177 tcgcgatcta gattataaaa agcaaacagg gccaacgtcc acacc 45
 180 <210> SEQ ID NO: 17
 181 <211> LENGTH: 6
 182 <212> TYPE: DNA
 183 <213> ORGANISM: Synthetic Construct
 185 <400> SEQUENCE: 17
 186 gctagc 6
 189 <210> SEQ ID NO: 18
 190 <211> LENGTH: 6
 191 <212> TYPE: DNA
 192 <213> ORGANISM: Synthetic Construct
 194 <400> SEQUENCE: 18
 195 aagctt 6
 198 <210> SEQ ID NO: 19
 199 <211> LENGTH: 6
 200 <212> TYPE: DNA
 201 <213> ORGANISM: Synthetic Construct
 203 <400> SEQUENCE: 19
 204 tctaga 6
 207 <210> SEQ ID NO: 20
 208 <211> LENGTH: 6
 209 <212> TYPE: DNA
 210 <213> ORGANISM: Synthetic Construct
 212 <400> SEQUENCE: 20
 213 tcgcga 6

Some errors

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/023,831

DATE: 04/16/2002

TIME: 14:35:38

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J023831.raw

216 <210> SEQ ID NO: 21
 217 <211> LENGTH: 21
 218 <212> TYPE: PRT
 219 <213> ORGANISM: Synthetic Construct
 221 <400> SEQUENCE: 21
 223 Asp Tyr Lys Asp Asp Asp Asp Lys Ala Ser Lys Leu Gly Ala Pro Gly
 224 1 5 10 15
 227 Pro Leu Gly Ile Ala
 228 20
 231 <210> SEQ ID NO: 22
 232 <211> LENGTH: 9
 233 <212> TYPE: PRT
 234 <213> ORGANISM: Synthetic Construct
 236 <400> SEQUENCE: 22
 238 Gly Pro Pro Gly Pro Pro Gly Leu Ala
 239 1 5
 242 <210> SEQ ID NO: 23
 243 <211> LENGTH: 9
 244 <212> TYPE: PRT
 245 <213> ORGANISM: Synthetic Construct
 247 <400> SEQUENCE: 23
 249 Gly Pro Pro Gly Pro Pro Gly Glu Arg
 250 1 5
 253 <210> SEQ ID NO: 24
 254 <211> LENGTH: 9
 255 <212> TYPE: PRT
 256 <213> ORGANISM: Synthetic Construct
 258 <400> SEQUENCE: 24
 260 Gly Pro Pro Gly Pro Pro Gly Pro Ala
 261 1 5
 264 <210> SEQ ID NO: 25
 265 <211> LENGTH: 9
 266 <212> TYPE: PRT
 267 <213> ORGANISM: Synthetic Construct
 269 <400> SEQUENCE: 25
 271 Gly Pro Pro Gly Pro Pro Gly Ala Pro
 272 1 5
 275 <210> SEQ ID NO: 26
 276 <211> LENGTH: 9
 277 <212> TYPE: PRT
 278 <213> ORGANISM: Synthetic Construct
 280 <400> SEQUENCE: 26
 282 Gly Leu Ala Gly Ala Pro Gly Leu Arg
 283 1 5
 286 <210> SEQ ID NO: 27
 287 <211> LENGTH: 45
 288 <212> TYPE: PRT
 289 <213> ORGANISM: Synthetic Construct
 291 <400> SEQUENCE: 27

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/023,831

DATE: 04/16/2002

TIME: 14:35:38

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J023831.raw

```

293 Gly Gly Lys Gly Asp Ala Gly Ala Pro Gly Glu Arg Gly Pro Pro Gly
294 1          5          10          15
297 Leu Ala Gly Ala Pro Gly Leu Arg Gly Gly Ala Gly Pro Pro Gly Pro
298          20          25          30
301 Glu Gly Gly Lys Gly Ala Ala Gly Pro Pro Gly Pro Pro
302          35          40          45
305 <210> SEQ ID NO: 28
306 <211> LENGTH: 7
307 <212> TYPE: PRT
308 <213> ORGANISM: Synthetic Construct
310 <400> SEQUENCE: 28
312 Met Gly Ala Pro Gly Ala Pro
313 1          5
316 <210> SEQ ID NO: 29
317 <211> LENGTH: 6
318 <212> TYPE: PRT
319 <213> ORGANISM: Synthetic Construct
321 <400> SEQUENCE: 29
323 Gly Ala Pro Gly Ala Pro
324 1          5
327 <210> SEQ ID NO: 30
328 <211> LENGTH: 7
329 <212> TYPE: PRT
330 <213> ORGANISM: Synthetic Construct
332 <400> SEQUENCE: 30
334 Gly Leu Glu Gly Pro Arg Gly
335 1          5
338 <210> SEQ ID NO: 31
339 <211> LENGTH: 26
340 <212> TYPE: DNA
341 <213> ORGANISM: Synthetic Construct
343 <400> SEQUENCE: 31
344 aattccatgg gtgctccagg tgctcc                26
347 <210> SEQ ID NO: 32
348 <211> LENGTH: 22
349 <212> TYPE: DNA
350 <213> ORGANISM: Synthetic Construct
352 <400> SEQUENCE: 32
353 ggccacctgg tggacctggt gg                22
356 <210> SEQ ID NO: 33
357 <211> LENGTH: 22
358 <212> TYPE: DNA
359 <213> ORGANISM: Synthetic Construct
361 <400> SEQUENCE: 33
362 ggcccgggtg taagggtgac gc                22
365 <210> SEQ ID NO: 34
366 <211> LENGTH: 20
367 <212> TYPE: DNA
368 <213> ORGANISM: Synthetic Construct

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/023,831

DATE: 04/16/2002

TIME: 14:35:39

Input Set : A:\EP.txt

Output Set: N:\CRF3\04162002\J023831.raw